ANN for hybrid modelling of batch and fed-batch chemical reactors

Yessin Ammar, Patrick Cognet, Michel Cabassud *

Laboratoire de Génie Chimique, Université de Toulouse, CNRS, INPT, UPS, Toulouse, France

HIGHLIGHTS

- Unconventional modelling based on ANN to rapidly develop a model from batch experiments.
- Recurrent ANN's (one per species) are assembled to predict time evolution of concentrations.
- Balanced esterification reaction of methanol by acetic acid is chosen as application.
- The global neural model is integrated in a reactor hybrid model.
- The hybrid model permits to transpose the reaction to a semi batch chemical reactor.

ARTICLE INFO

Keywords:
Artificial neural networks
Hybrid model
Kinetics modelling
Chemical reactor
Esterification

ABSTRACT

An unconventional modelling methodology based on artificial neural networks is proposed to rapidly develop a model from data obtained during different batch experiments.

The objective of the global model is to predict time evolution of concentrations of all species present in the reaction medium. For this, different recurrent neural networks are elaborated to estimate a particular species as a function of operating parameters and concentrations of all species and then assembled in a complex global model.

To validate the approach, the esterification reaction of methanol by acetic acid, which presents equilibrium, has been chosen. The kinetic evolution of the chemical species during experiments conducted in batch mode are correctly represented whatever the operating conditions. Finally, the global model based on neural networks is integrated in a hybrid model. This permits to transpose the reaction to a semi batch chemical reactor which has not been considered during the learning phase.

1. Introduction

In the field of fine chemicals or pharmaceuticals, the main apparatus for the implementation of chemical syntheses still remains the batch reactor. An alternative to using this type of reactor is to transpose the reactions in plug flow continuous reactors with intensification of the chemical synthesis (Anxionnaz et al., 2010). The interest comes from better control of heat exchange and the possibility of contacting the reactants in an optimal way to achieve better performance and greater selectivity. This transition of the chemical industry to process intensification (Stankiewicz and Moulijn, 2002) will lead to implement the reactions in a different way from that currently practiced favouring concentrated media, high temperature and the use of catalysts. Under these conditions, simple kinetic laws of Arrhenius type are no longer sufficient to properly model the evolution of the medium. This trend towards process intensification will require the contribution of numerical simulation tools and thus the development of specific methodologies to rapidly establish kinetic models used for simulation, optimization or control. Indeed, for safety, economic efficiency and environmental protection, understanding and predicting the behaviour of a chemical reaction are of great importance in industry.

In literature, modelling approaches by models of representation have been proposed. The basic assumption is that the reaction rate usually follows an Arrhenius type law, that is to say exponentially activated by temperature and influenced by the concentrations of the reactants according to orders to be determined. While this approach has yielded interesting results in the case of reactions carried out in diluted solvent medium (Cabassud et al., 2005), it is noticeable that when the mode of production is changed and moved to an intensified one with increased concentrations of reagents and introduction of a catalyst, such an approach is no longer sufficient to properly model the evolution of the medium (Balland et al., 2002).

* Corresponding author.
E-mail address: michel.cabassud@ensiacet.fr (M. Cabassud).

https://doi.org/10.1016/j.ces.2021.116522

Nomenclature

Symbols

C	concentration (mol. m ^{3})
Ea	activation energy (J mol ^{1})
F	molar flow rate (mol s ^{1})
k0	pre exponential factor (m^3 mol ^{1} s ^{1} mL$_{catalyst}^{1}$)
M	molecular weight (kg mol ^{1})
N	number of moles (mol)q ^{1}
Q	volumetric flow rate (m^3 s ^{1})
Q_{cat}	amount of catalyst (mL)
r	reaction rate (mol s ^{1} m ^{3})
R	perfect gas constant (=8.314 J K ^{1} mol)
t	time (s)
T	Temperature (K)
V	Volume (m^3)

Greek letters

A	stoichiometric coefficient
ρ	density (kg m ^{3})
Δt	Sampling time (s)

Subscripts

AA	Acetic Acid
Cat	Catalyst
E	esterification
H	hydrolysis
M	Methanol
MA	Methyl Acetate
W	Water

The rapid development of on line analytical techniques (FTIR, NIR, Raman Spectroscopy ...) suggests the possibility to dispose of a big number of data allowing characterizing the evolution of the concentrations of the different species during the implementa tion of a chemical reaction. Indeed, traditionally, the reaction kinetics are determined from the measurements obtained by external analysis (e.g. by chromatography) carried out on samples taken at different reaction times. This methodology faces numer ous problems (sampling and sample storage, sample representa tiveness ...). Therefore, in recent years in situ spectroscopic measurements have been developed that provide many benefits. Measurements by UV/VIS wavelengths are the most common tech niques used but other spectroscopic techniques such as near infra red (NIR), infrared (IR), Raman and multi wavelength ultraviolet/ visible (UV/VIS) have also been used for kinetic studies (Bayada et al., 1995; Bijlsma et al., 1998; Dyson et al., 2000; Furusjo et al., 1998, 2000a, 2000b, Furusjö et al., 2003; Molloy et al., 1999; Mozharov et al., 2011; Puxty et al., 2005, 2008, Rizkin et al., 2019).

These methods have proven their efficiency for both qualitative and quantitative analysis in different fields. They are fast and mainly non destructive compared to the medium to analyse. The recent development of instrumentation, especially optical fibers, which can deport the measurement, has increased the interest in these methods.

There, it is the field of "Chemometrics" in full development, which implements mathematical methods and statistical treat ment methods (modelling, classification, principal component analysis ...) to select optimal experimental procedures and pro cessing of chemical analysis. Thus, after treatment, these on line analysis methods allow to obtain the evolution of concentrations of all chemical species present in the reaction medium with acqui sition times close together, mainly limited by the data processing time.

In this paper, an unconventional modelling is proposed by using a technique that does not require explicit knowledge of physical laws (as it will be extracted from the data) but requires data in number. Neural networks, used successfully in the modelling of reactors and chemical reactions (Hosen et al., 2014; Molga, 1983; Mujtaba et al., 2006; Sunphorka et al., 2017), offer the advantage of being a nonlinear multivariable tool, configurable by learning on experimental representative data. Moreover, the development of a hybrid model gives the possibility to include all or part of the available physical knowledge.

In order to generalize the developed methodology to a large number of chemical syntheses and without a priori knowledge of the reaction scheme and therefore the number of reactions to be implemented, it has been decided to work on the direct prediction of concentrations of all species, rather than on rates or extents of the reactions. Thus, the objective of the global model that will be developed is to predict time evolution of concentrations of all spe cies present in the reaction medium. For this, recurrent neural net works are implemented, whose delay is equivalent to a sampling period.

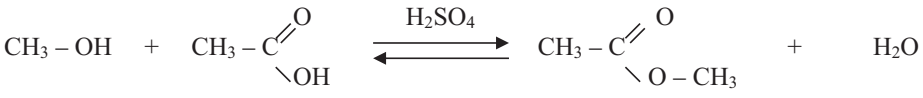
If only one network is considered for the determination of all the concentrations, it faces a problem of convergence. Therefore, the development of a much more complex model which associates as many networks as there are species has been preferred. Each neural network is elaborated to estimate a particular species as a function of operating parameters and concentrations of other species.

To validate the approach, the esterification reaction of methanol by acetic acid has been chosen as application. Then, it is shown that with such a model based on neural networks it is possible to cor rectly represent the kinetic evolution of the chemical species dur ing experiments conducted in batch mode whatever the operating conditions. Finally, the most interesting results are obtained after wards with the integration of the neural model in a hybrid model that permits to represent the operation of a semi batch chemical reactor which has not been considered during the learning phase of the different neural networks.

2. Characterization of the esterification reaction

2.1. Chemical reaction

To illustrate the approach proposed in this paper, the Fisher esterification reaction between a carboxylic acid and an alcohol to form an ester and water has been chosen. This reaction is rever sible and equilibrated. It is performed in homogeneous phase and catalysed by sulfuric acid and leads to the formation of methyl acetate (MA) and water (W) from acetic acid (AA) and methanol (M) according to the following equation:



(1)

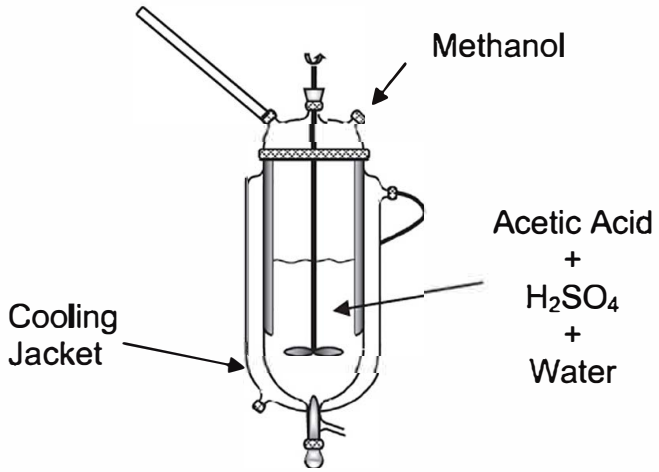


Fig. 1. Scheme of the experimental reactor.

Experiments have been performed in a batch glass reactor of an effective capacity of 1.5 L (see Fig. 1). The reactor is equipped with a jacket that allows temperature control. During the experiments, methanol is fed into a mixture of dilute acetic acid and sulfuric acid (catalyst). Water, a product of the reaction, is also present in the initial mixture.

2.2. Kinetic model and parameters

Kinetic modelling of the esterification of acetic acid in the presence of methanol has been the subject of numerous studies. A simple model based on an Arrhenius formulation from the works of Smith (1939) and Agreda et al. (1990) has been proposed. The expression of the reaction rate associated with this model is given by the following equation:

$$r \quad k0_e Q_{cat} \exp\left(\frac{Ea_e}{RT}\right) C_{AA} C_M \quad k0_h Q_{cat} \exp\left(\frac{Ea_h}{RT}\right) C_{MA} C_W \qquad (2)$$

With:

$k0_e$: Pre exponential factor of the esterification reaction ($m^3\,mol^{-1}\,s^{-1}\,mL_{cat}^{-1}$)
$k0_h$: Pre exponential factor of the hydrolysis reaction ($m^3\,mol^{-1}\,s^{-1}\,mL_{cat}^{-1}$)
Q_{cat}: Amount of catalyst (mL)
Ea_e: Activation energy of the esterification reaction (J mol^{-1})
Ea_h: Activation energy of the hydrolysis reaction (J mol^{-1})
T: Temperature (K)
R: perfect gas constant (=8.314 J K^{-1} mol)
C_{AA}: Acetic Acid concentration (mol. m^{-3})
C_M: Methanol concentration (mol. m^{-3})
C_{MA}: Methyl Acetate concentration (mol. m^{-3})
C_W: Water concentration (mol. m^{-3})

For the present work, values of kinetic parameters ($k0_1$, $k0_2$, Ea_1, Ea_2) estimated in a previous work from tests carried out in batch reactor (Elgue et al., 2009) have been considered.

2.3. Batch reactor model

To simulate the evolution of the different components in batch mode, a model has been developed from the resolution of mass balances within the reactor. The equations are given by:

$$\frac{dN_i(t)}{dt} \quad \alpha_i V r(t) \qquad (3)$$

with:

$$r(t) \quad f(T, Q_{cat}, C_i(t)) \qquad (4)$$

$$C_i(t) \quad \frac{N_i(t)}{V} \qquad (5)$$

With:

$N_i(t)$: Number of moles of component (i) at time t (mol)
$C_i(t)$: concentration of component (i) at time t (mol m^{-3})
V: Volume of the reaction mixture (m^3)
$\alpha_i = 1$ for methyl acetate and water
$\alpha_i = -1$ for acetic acid and methanol
t: time (s)
r (t): reaction rate at time t, function of temperature, catalyst amount and components concentration (mol s^{-1} m^{-3})

This model allows simulating the esterification reaction under different operating conditions and for different initial concentrations of acetic acid, methanol and water. The model was validated on experimental data in a temperature range between 50 and 70 °C and an amount of catalyst from 5 to 10 mL (Elgue et al., 2009).

In the following part of this article, this model is used to generate the data needed for the constitution of the databases for learning, testing and validation. In this way, it is possible to simulate the experimental data that would be obtained with an on line measuring system like FTIR, NIR or Raman spectroscopy which provides a very large number of data with a given sampling period. In this paper, the sampling period of the generated data is fixed to 2 s.

3. Artificial neural networks (ANN)

As far as a large number of data can be obtained from pilot plants, statistical models appear as promising tools for developing simulation models. Among such tools, the model chosen in this work lies on a combination of artificial neural networks (ANNs).

ANNs, as any models can be used to model relationships between input and output variables. Their main interest is that ANNs are statistical and nonlinear tools. They are interesting tools to model complex phenomena and processes for which data are numerous and noisy. Their structure is obtained by learning from a set of experimental data which constitute a learning database. They can also be associated to available laws based on knowledge of the process.

In process engineering, artificial neural networks have so far mainly been used in process modelling, process control, fault diagnosis, error detection, data reconciliation and process analysis (Funashi, 1989; Bulsari, 1995; Dirion et al., 1995; Fakhr Eddine and Cabassud, 1996; Delgrange Vincent et al., 2000; Chouai et al., 2000a,b).

One important aspect in an ANN development procedure is the learning process. Representative examples are presented to the network so that it can integrate the corresponding knowledge within its structure. The learning process consists in determining the values of the model parameters called weights ($w_{j,i,k}$) that produce the best fit between the actual and the predicted outputs over the entire training region.

The topology of the neural network determines the accuracy and the degree of representation of the model. A number of papers have shown that a feedforward network has the potential to approximate any non linear function. In this paper, only one hidden layer has been considered. The number of neurons in this hidden layer has been chosen by trial and error tests. Many different network architectures are used; the most popular of these is the multilayer network with sigmoid activation functions for the learning phase. Theoretical and numerical results proved that

quasi Newton algorithms are superior to steepest algorithms. Watrous (1987) employed and compared DFP and BFGS methods with the backpropagation algorithm: this comparison showed that DFP and BFGS need fewer iterations. For this reason, a quasi Newton (Dennis and Schnabel, 1983) leaming algorithm has been used to train the different neural networks.

4. Modelling of the esterification reaction by neural networks

4.1. Structure of the model

In the case of chemical transformations, the literature proposes different types of approach using the concentrations of reagents and products or reaction extent (Molga, 1983). In the case where it is assumed that stoichiometry is not known, it seemed more appropriate to consider only the concentrations of chemical spe cies. (See Table1).

The structure developed (Fig. 2) to estimate the evolution of species during the esterification of methanol by acetic acid from the initial operating conditions includes 6 inputs: Temperature (T), amount of catalyst (Q_{cat}) and respective concentrations of acetic acid (C_{AA}), methanol (C_M), methyl acetate (C_{MA}) and water (C_W). This structure (Fig. 2) must make it possible to determine the four outputs that are the concentrations of acetic acid, metha nol, methyl acetate and water obtained after a time equivalent to the sampling period.

The adopted structure consists in breaking down the four calcu lations and to construct four different neural networks (see Table 2): each one is devoted to calculate the concentration of a single component (reactant or product). Different network struc tures, with different number of neurons in the hidden layer, have been tested. For each network R1, R2, R3 and R4, the one leading to the lowest error on the test database is retained.

For each network, leaming is carried out separately. Four differ ent sets of weights corresponding to the four networks (R1, R2, R3, R4) are then obtained by this way. The four networks are then interconnected as shown in the block diagram on Fig. 3. The result ing model, called 4ANN, works recursively. It is then tested on its ability to reproduce a complete dynamic experiment.

4.2. Elaboration of databases

The model (4ANN) described above must be capable of simulat ing the esterification reaction according to several operating condi tions and for different initial concentrations. The temperature varies between 50 and 80° C and the amount of catalyst from 5 to 10 mL for a volume of 1 L.

Depending on the chosen operating conditions, the ranges of variation given in Table 3 have been considered:

Simulations performed using the kinetic model presented in II.2 have allowed to generate databases. These databases are then used for training and testing the different neural networks that are developed.

To cover the operating range while limiting the number of runs considered during learning, experimental design has been used to plan the number and the operating conditions of these runs (Mathieu and Phau Tan Luu, 1981).

Table 1
Kinetic parameters of the esterification reaction.

$k0_e$	$k0_h$	Ea_e	Ea_h
$m^3 \cdot mol^{-1} \cdot s^{-1} \cdot mL_{cat}^{-1}$		$J \cdot mol^{-1}$	
4.21	0.322	53,804	52,584

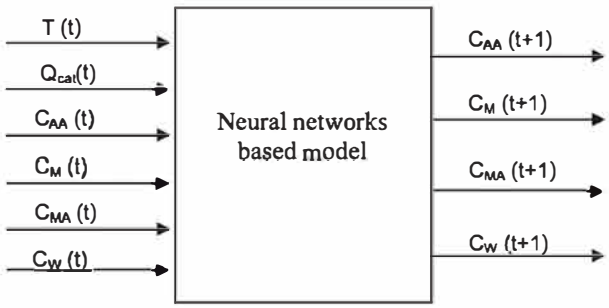


Fig. 2. Global objective of the developed neural networks structure.

4.2.1. Matrix of experiments

Experimental design allowed to define the operating conditions. The amount of methyl acetate is initially equal to zero, and conse quently it is not a factor for the experimental design. So this latter can be reduced to four main factors: Temperature (°C), amount of catalyst (ml), amount of acid (in mol) and amount of water (in mol). The amount of methanol (mol) to be added is calculated to maintain an initial constant volume of one litre.

With four factors and a spherical domain, the matrix of experi ments then includes 16 runs.

4.2.2. Learning data base

For the leaming database, factors X1, X2, X3 and X4 correspond respectively to the amount of acetic acid (Q_{AA} in mol), water (Q_w in mol), sulfuric acid catalyst (Q_{cat} in mL) and temperature (T in °C). Using variations of four factors presented in the Table 3 and calcu lating the amount of methanol (Q_M in mol) to adjust a reaction vol ume equal to one litre, the following matrix is obtained (Table 4):

4.2.3. Test database

For the test database, the same factors as for the learning data base are considered and the same calculations are applied to obtain the initial composition in methanol. The retained operating condi tions are summarized in Table 5:

The field of application is restricted by the minimum and the maximum of every variable, and corresponds to the domain of validity of the neural model. For every parameter, an interval has been defined according to the observed values. These min and max values given in Table 6 are used to normalize the raw values between 0.1 and 0.9.

4.3. Sampling data and number of data

The sampling period was chosen equal to 2 s for data acquisition.

Yet, the number of points considered for each learning curve has a great influence on the development of the network and the optimization of its result. For efficient learning, this number must be sufficient to allow the network to detect all the variation of the curve. However, as the esterification reaction is balanced, if one considers too long reaction times, there is a risk of "over leaming" of the part corresponding to the equilibrium at the expense of the purely reaction phase. Learning from reaction time of 2000 s, led to networks that allowed the data to be represented in the most accurate way.

Moreover, during the learning phase, the acquisition time was kept to Δt = 2 s but only one point out of 5 was considered to lighten the calculation. Then, 200 points were used to cover an esterification reaction that lasts for 2000 s.

Table 2
Inputs and outputs of the four neural networks.

	R1	R2	R3	R4
Inputs	$T(t)$, $Q_{cat}(t)$, $C_{AA}(t)$, $C_M(t)$, $C_{MA}(t)$, $C_W(t)$	$T(t)$, $Q_{cat}(t)$, $C_{AA}(t)$, $C_M(t)$, $C_{MA}(t)$, $C_W(t)$	$T(t)$, $Q_{cat}(t)$, $C_{AA}(t)$, $C_M(t)$, $C_{MA}(t)$, $C_W(t)$	$T(t)$, $Q_{cat}(t)$, $C_{AA}(t)$, $C_M(t)$, $C_{MA}(t)$, $C_W(t)$
Output	$C_{AA}(t+1)$	$C_M(t+1)$	$C_{MA}(t+1)$	$C_W(t+1)$

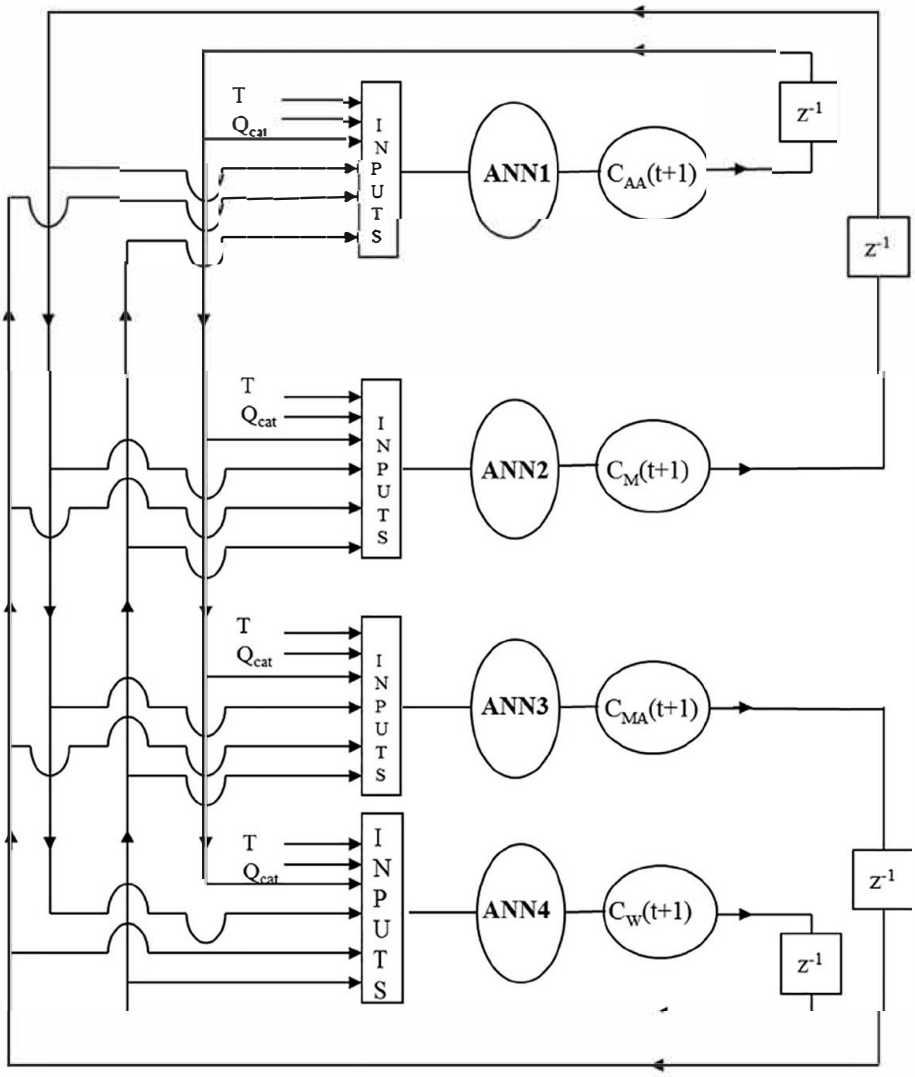


Fig. 3. Global architecture of the developed model (4ANN) built from four specific neural networks.

Table 3
Range of variation of the inputs for the developed neural networks.

T [°C]		Q_{cat} [mL]		C_{AA} [mol·L^{-1}]		C_M [mol·L^{-1}]		C_W [mol·L^{-1}]	
Min	Max	Min	Max	Min	Max	Min	Max	Min	Max
50	80	5	10	8	15	0.5	11	0	12

4.4. Number of neurons in the hidden layer

Neurons in the hidden layer can increase the number of internal parameters of the network, and therefore the non linear model. With the properties of the function f used (sigmoid function), the use of a single hidden layer is sufficient to approach any non linear function (Rivals, 1995). Therefore, the networks used in this work consist of three layers.

A study has been conducted to determine the number of neu rons in the hidden layer representing the best compromise between data representation and generalization ability of the resulting model.

Table 4
Initial conditions of the runs used for the training set.

Run	Q_{AA} [mol]	Q_M [mol]	Q_W [mol]	Q_{Cat} [mL]	T [°C]
Ldb1	11.50	5.77	6,00	7.50	80.00
Ldb2	9.19	10.79	2.04	5.85	66.55
Ldb3	13.80	4.26	2.04	5.85	66.55
Ldb4	9.19	7.27	9.95	5.85	66.55
Ldb5	13.80	0.75	9.95	5.85	66.55
Ldb6	9.19	10.79	2.04	9.14	66.55
Ldb7	13.80	4.26	2.04	9.14	66.55
Ldb8	9.19	7.27	9.95	9.14	66.55
Ldb9	13.80	0.75	9.95	9.14	66.55
Ldb10	8.00	10.72	6.00	7.50	48.62
Ldb11	15.00	0.82	6.00	7.50	48.62
Ldb12	11.50	8.44	0.00	7.50	48.62
Ldb13	11.50	3.10	12.00	7.50	48.62
Ldb14	11.50	5.77	6.00	5.00	48.62
Ldb15	11.50	5.77	6.00	10.00	48.62
Ldb16	11.50	5.77	6.00	7.50	60.00

Table 5
Initial conditions of the runs used for the test database.

Run	Q_{AA} [mol]	Q_M [mol]	Q_W [mol]	Q_{Cat} [mL]	T [°C]
Tdb1	11.50	5.77	6.00	10.00	65.00
Tdb2	9.19	10.79	2.04	7.98	55.11
Tdb3	9.19	10.79	2.04	7.98	74.88
Tdb4	13.80	4.26	2.04	7.98	55.11
Tdb5	13.80	4.26	2.04	7.98	74.88
Tdb6	9.19	7.27	9.95	7.98	55.11
Tdb7	9.19	7.27	9.95	7.98	74.88
Tdb8	13.80	0.75	9.95	7.98	55.11
Tdb9	13.80	0.75	9.95	7.98	74.88
Tdb10	11.50	5.77	6.00	5.29	50.00
Tdb11	11.50	5.77	6.00	5.29	80.00
Tdb12	8.00	10.72	6.00	5.29	65.00
Tdb13	15.00	0.82	6.00	5.29	65.00
Tdb14	11.50	8.44	0.00	5.29	65.00
Tdb15	11.50	3.10	12.00	5.29	65.00
Tdb16	11.50	5.77	6.00	7.00	65.00

Table 6
Values used to normalize the inputs and outputs of the different neural networks.

T [K]		Q_{cat} [mL]		C_{AA} [mol·L ^{1}]		C_M [mol·L ^{1}]		C_{MA} [mol·L ^{1}]		C_W [mol·L ^{1}]	
Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max
273	373	0	15	0	20	0	20	0	20	0	20

Table 7
Error on the learning data base and the test data base based on the number of hidden neurons for the four different neural networks.

	Acetic Acid		Methanol		Methyl Acetate		Water	
Number of neurons	Errror on the learning data base	Error on the test data base	Errror on the learning data base	Error on the test data base	Errror on the learning data base	Error on the test data base	Errror on the learning data base	Error on the test data base
6	3.6710 ^{9}	5.3610 ^{11}	7.8410 ^{10}	1.5610 ^{10}	2.2910 ^{9}	4.5910 ^{10}	4.5210 ^{9}	2.8810 ^{10}
8	2.5010 ^{9}	9.8710 ^{12}	8.0810 ^{10}	2.1710 ^{11}	2.7610 ^{9}	6.1810 ^{11}	5.8410 ^{9}	1.6210 ^{10}

Different initializations of biases and weights were tested with random initial values. All biases and weights leading to the small est value of the error calculated on the test database was retained. Table 7 presents the errors obtained with the final weights used for the simulations.

The error is computed as follow:

$$\text{Error} \quad \frac{1}{2N}\sum_k\sum_j\left[\widehat{C_{k,j}} \quad C_{k,j}\right]^2 \tag{6}$$

where k refers to the run considered in the data base and j to the sample point of this run (as said before j varies between 1 and 200). $\widehat{C}_{k,j}$ is the predicted ANN model output and $C_{k,j}$ is the reference output value.

It can be observed that, in most cases, errors are improved from 6 hidden neurons to 8, both on the learning data base and on the test one. However, the improvement is not very important and the levels reached are very satisfactory, which led to consider this number of eight hidden neurons for the rest of the study.

Once the four networks formed, they have been combined in a model built according to the architecture proposed on Fig. 3. The model works in a recurrent way from the initial conditions and cal

culates the concentrations of the different species every 2 s (time corresponding to the sampling period).

This model (4ANN) allowed the simulation of all the runs con stituting the learning and testing data bases. On the set of curves, a very good agreement between the initial curves and those obtained using the model 4ANN was observed, which confirms the efficiency of the elaborated model. These data are not pre sented in this paper, it has been preferred to focus on data not seen in the learning and test databases, as shown in the next section.

4.5. Validation of the (4ANN) model

Afterward and to validate the 4ANN model's ability to general ize, results obtained in conditions not seen during learning (con centrations, temperature and amount of catalyst) are presented (see Table 8). It is to be noticed that these curves do not belong either to the learning database or to the test database.

For the shake of clarity, the figures are presented in the follow ing form: data obtained by the 4ANN model correspond to the solid lines while data obtained experimentally (Figs. 7 and 8) or by the conventional kinetic model (Figs. 4 6) presented in part II.2 are symbolized by the points.

In a first time, the performance of the 4ANN model can be eval uated in terms of mean square error (MSE) criterion, which is defined by the following equation:

$$MSE \quad \frac{1}{4N}\sum_{l\ 1}^{4}\sum_{i\ 1}^{N}\left[\widehat{C_{l,j}} \quad C_{l,j}\right]^2 \qquad (7)$$

where:

l refers to the chemical species (Acetic Acid, Methanol, Methyl Acetate and Water)
j refers to the sample (as said in part IV.3, the sampling time has been chosen equal to 2 s, so N is equal to 1000 for experiments that last 2000 s)
$\widehat{C}_{l,j}$ is the predicted 4ANN model output and $C_{l,j}$ is the reference output value.

Low values of MSE are obtained in all cases which confirms the ability of the model to accurately describe changes in concentra tions during the different runs. To confirm this, three runs (Val2, Val3 and Val4) are presented in detail. Figs. 4 6 show that the model (4ANN) integrating the four neural networks reproduces perfectly the evolution of the different components (reactants and products) during the progress of the reaction for very different operating conditions given in Table 8.

4.6. Validation on experiments

To confirm the 4ANN model quality, its validation has also been conducted on a data base experimentally established in the labora tory according to the protocol given Section 2.1. Two experiments are shown on Figs. 7 and 8 which correspond to the following oper ating conditions (Table 9):

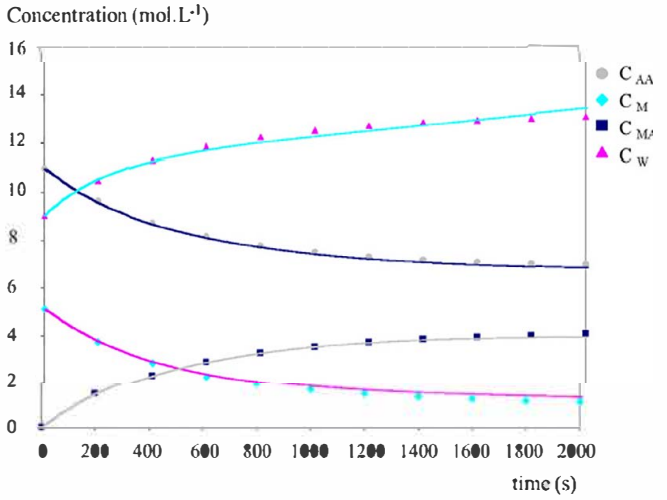


Fig. 4. Comparison of the time evolution of the different components predicted by the 4ANN model (solid line) and the conventional kinetic model (points) – Esterification at 70 °C with 6 mL of catalyst (Val 2).

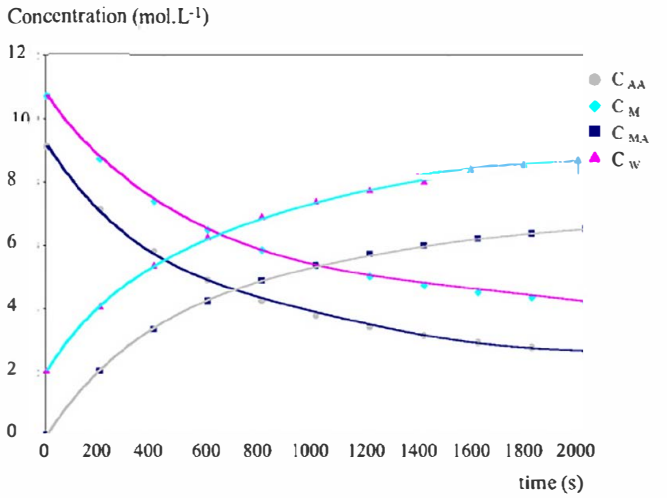


Fig. 5. Comparison of the time evolution of the different components predicted by the 4ANN model (solid line) and the conventional kinetic model (points) – Esterification at 66.55 °C with 5.85 mL of catalyst (Val 3).

For these operating conditions, the results demonstrate the effi ciency of the 4ANN developed model. A good fitting of the evolu tion of the concentrations of the compounds obtained experimentally is ensured by the 4ANN model.

Globally, the different studies which have been carried out so far, give interesting and promising results. After learning, the model developed makes it possible to correctly represent the evo lution of the esterification reaction of methanol by acetic acid under different operating conditions, in particular those which have not been used during the different phases of learning.

Table 8
Initial conditions of runs used for validation.

Run	Q_{AA} [mol·L ^{1}]	Q_M [mol·L ^{1}]	Q_W [mol·L ^{1}]	Q_{Cat} [mL]	T [°C]	MSE
Val 1	12.00	4.621	7.00	7.00	65.00	0.0107
Val 2 – Fig. 4	11.00	5.154	9.00	6.00	70.00	0.0154
Val 3 – Fig. 5	9.19	10.79	2.04	5.85	66.55	0.0024
Val 4 – Fig. 6	10.00	5.67	11.00	5.00	50.00	0.0396
Val 5	12.50	5.24	4.00	5.00	55.00	0.0660

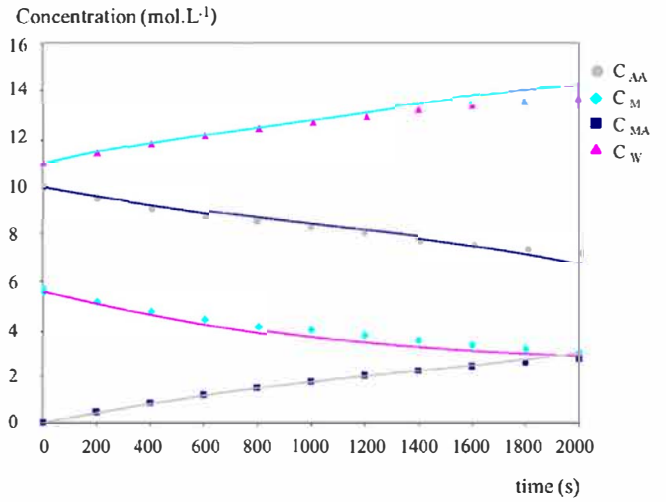


Fig. 6. Comparison of the time evolution of the different components predicted by the 4ANN model (solid line) and the conventional kinetic model (points) – Esterification at 50 °C with 5 mL of catalyst (Val 4).

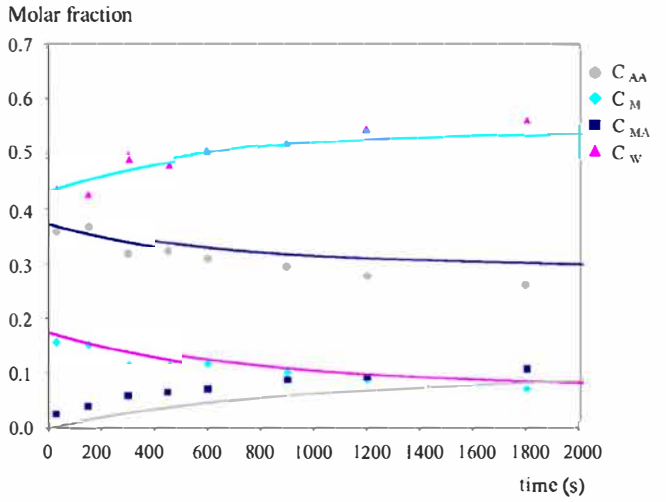


Fig. 7. Comparison of the time evolution of the different components predicted by the 4ANN model (solid line) and experimental data (points) – Esterification at 48 °C with 10 mL of catalyst (Exp 1).

At this stage, the 4ANN developed model, which couples four recurrent neural networks, and the architecture of which is presented on Fig. 4, answers very well the objective which is to propose a methodology for the rapid development of a stoechiokinetic model of a chemical transformation.

5. Application of the 4ANN model to a fed-batch reactor

In this section, the integration of the 4ANN model in a hybrid structure is discussed. The goal is to show how it is possible to use the 4ANN model to transpose the operating mode of the reaction from a batch to a fed batch operation with reactant feeding.

5.1. Fed batch reactor

The semi batch or fed batch mode is highly used in the industry, especially in the fine chemicals operations. Some components are added to the load during the reaction. This type of operating mode is mainly used when the reactions are fast and highly exothermic. Using a semi continuous operation mode allows to

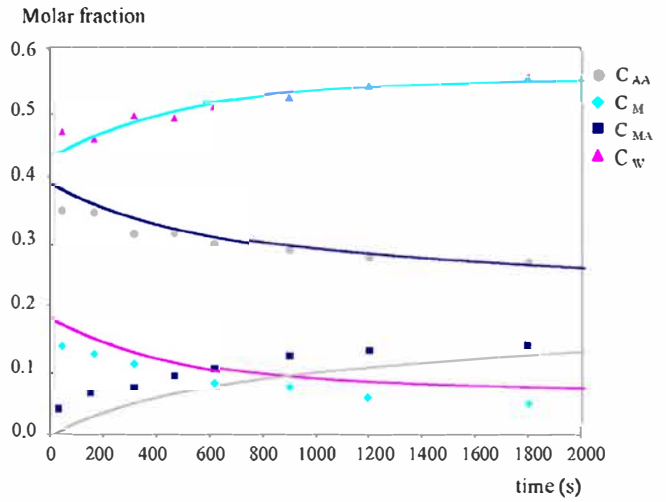


Fig. 8. Comparison of the time evolution of the different components predicted by the 4ANN model (solid line) and experimental data (points) – Esterification at 57 °C with 10 mL of catalyst (Exp 2).

control the rate of heat release, thereby maintaining the thermal regime within the limits.

To validate the results of simulations obtained by integrating the 4ANN neural network model in a more general hybrid structure, a reference database is needed. This latter has been generated with a conventional model of semi continuous reactor using the Arrhenius type kinetic model presented in paragraph II.2 (Table1).

5.2. Modelling of esterification reaction in a fed batch reactor

Initially the load is composed of the first reagent (acetic acid), water and the catalyst (sulfuric acid). The operation mode follows two main stages: firstly, the addition of the second reagent (Methanol) at a fixed flow rate and secondly, a time period after the end of feeding to reach equilibrium. It is assumed that the reactor is perfectly controlled from a thermal point of view and therefore the temperature remains constant.

The modelled system is a dynamic semi continuous process. The mathematical model considers the differential terms characteristics of the dynamic behaviour of the process, which is the set of ordinary differential equations (ODE). Then, the program which solves the ODE system allows to predict the dynamic behaviour of the system and to determine changes in the concentrations and volume over time.

The first calculation step corresponds to the addition of methanol to the reactor, and the system of differential equations is as follows:

$$\frac{dN_{AA}(t)}{dt} \quad V(t)r(t) \tag{8}$$

$$\frac{dN_{M}(t)}{dt} \quad F_{M}(t) \quad V(t)r(t) \tag{9}$$

$$\frac{dN_{MA}(t)}{dt} \quad V(t)r(t) \tag{10}$$

$$\frac{dN_{W}(t)}{dt} \quad V(t)r(t) \tag{11}$$

$$\frac{dV(t)}{dt} \quad Q(t) \tag{12}$$

$$Q(t) \quad \frac{F_{M}(t)M_{M}}{\rho_{M}} \tag{13}$$

Table 9
Operating conditions for the experimental results presented on Figs. 10 and 11.

Experiment	Q_{AA} [mol·L ^{1}]	Q_M [mol·L ^{1}]	Q_W [mol·L ^{1}]	Q_{Cat} [mL]	T [°C]
Exp 1 – Fig. 7	10.25	4.89	11.91	10.00	48.00
Exp 2 – Fig. 8	10.61	4.91	10.74	10.00	57.00

With:

$N_{AA}(t)$: Moles number of Acetic Acid at time t (mol)
$N_M(t)$: Moles number of Methanol at time t (mol)
$N_{MA}(t)$: Moles number of Methyl Acetate at time t (mol)
$N_W(t)$: Moles number of Water at time t (mol)
$V(t)$: Volume at time t (m^3)
$F_M(t)$: Molar flow rate of Methanol at time t (mol s^{-1})
$Q(t)$: Total volumetric flow rate at time t (m^3 s^{-1})
r (t): reaction rate at time t (mol m^{-3} s^{-1})
M_M: molecular weight of Methanol (kg mol^{-1})
ρ_M: density of Methanol (kg m^{-3})

In the second calculation step, the addition of Methanol is stopped and therefore the molar flow rate of methanol and the volumetric flow rate are equal to zero.

This simulation program allows generating the data that are used later for checking the ability of the 4ANN model to be integrated in a hybrid structure in the case of a fed batch reactor.

Fig. 9 exemplifies the difference of time evolution of the different components' concentrations for the implementation of the esterification reaction in batch and fed batch mode. The feeding time is set to 200 s while the total amount of reactants is identical in both cases. Therefore, if the concentration profiles are different at the beginning of the operation, they finally meet at equilibrium and overlap.

5.3. Hybrid modelling including the 4ANN model

In paragraph IV, a model based on neural networks (called 4ANN) has been developed for modelling the esterification reaction of acetic acid by methanol in the presence of sulfuric acid as catalyst. In the following, it will be shown how this model (4ANN) can be integrated into a hybrid structure for the simulation of a fed batch reactor. This model must be able to consider the two stages of calculations: firstly, the addition of methanol to the reactor and secondly, after the end of addition, the evolution of the concentrations of the various components (reagents and products).

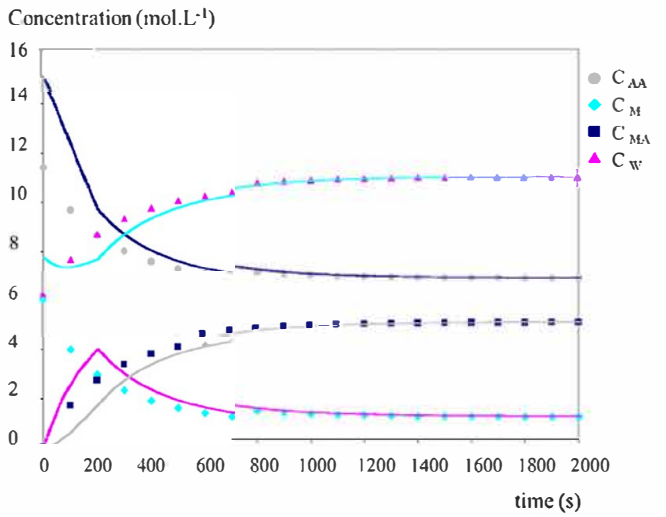


Fig. 9. Comparison of the components evolution during the esterification reaction according to the operating mode [batch mode (points) and semi-batch mode (solid line)] (Ldb1).

The equations expressing the material balance for a reagent during feeding are written as follows:

$$\frac{dN_i(t)}{dt} \quad \alpha_i V(t) r(t) + F_i(t) \tag{14}$$

$$\text{with} N_i(t) \quad C_i(t)V(t) \tag{15}$$

Using a first order approximation, Eq. (12) can be discretized as follows:

$$\frac{N_i(t+1) \quad N_i(t)}{\Delta t} \quad \alpha_i V(t) r(t) + F_i(t) \tag{16}$$

$$N_i(t+1) \quad C_i(t)V(t) + \alpha_i V(t) r(t)\Delta t + F_i(t)\Delta t \tag{17}$$

$$N_i(t+1) \quad (C_i(t) + \alpha_i r(t)\Delta t)V(t) + F_i(t)\Delta t \tag{18}$$

The term $(C_i(t) + \alpha_i\ r(t)\ \Delta t)$ is in fact computed by the 4ANN model and is called $C_{i0}(t+1)$ in the following.

To integrate the model based neural networks developed (4ANN) which takes into account the kinetic contribution, the following procedure has been established:

Step 1: Knowing concentrations $C_i(t)$ at time t, concentrations $C_{i0}(t+1)$ are computed at time $t + \Delta t$ using the 4ANN mode to estimate the contribution due to the reaction.

Step 2: Add the feeding term (moles):

$$N_i(t+1) \quad C_{i0}(t+1)V(t) + F_i(t)\Delta t$$

Step 3: Correction of the volume:

$$V(t+1) \quad V(t) + Q(t)\Delta t$$

Step 4: Computation of concentrations Ci(t + 1)

$$C_i(t+1) \quad N_i(t+1)/V(t+1)$$

With:

$N_i(t)$: Mole number of component i at time t (mol)
$N_{i0}(t)$:: Intermediary mole number of component i at time t (mol)
$C_i(t)$: Concentration of component (i) at time t (mol m^{-3})
$V(t)$: Volume of the reaction mixture at time t (m^3)
$F_i(t)$: Molar flow rate of component (i) at time t (mol s^{-1})
$Q(t)$: Total volumic flow rate at time t (m^3 s^{-1})
Δt: Sampling time (s)
r(t): Reaction rate at time t (mol s^{-1} m^{-3})
$\alpha(i)$: Stoichiometric coefficient of component I (1 for AA and M, +1 for MA and W)

This procedure results in the diagram presented on Fig. 10.

This hybrid model that integrates the 4ANN model obtained above has been used to simulate the behaviour of the reaction implemented in fed batch mode, without going back into learning for the four different neural networks.

The curves corresponding to different operating conditions listed in Table 10 are presented on Figs. 11 14.

Time evolution of the different component concentrations predicted by the hybrid model are compared to those determined by the conventional Arrhenius type kinetic model presented Sec

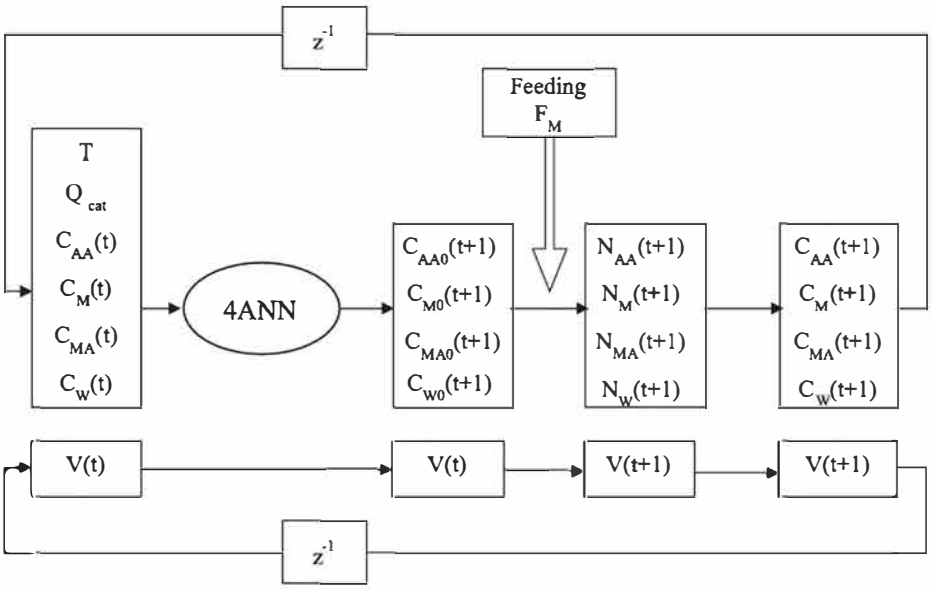


Fig. 10. Global architecture of the hybrid model for fed-batch representation.

Table 10
Operating conditions of the fed-batch runs.

Run	T [°C]	Q_{cat} [mL]	C_{AA}[mol·L ^{1}]	C_W[mol·L ^{1}]	feeding time [s]	F_M[mol·s ^{1}]	Number of moles of M fed [mol]
FB1 – Fig. 11	66.55	9.14	16.06	3.57	200	0.0539	10.79
FB2 – Fig. 12	80.00	7.50	14.86	7.75	200	0.0288	5.77
FB3 – Fig. 13	60.00	8.00	13.15	13.15	200	0.0305	6.11
FB4 – Fig. 14	60.00	8.00	13.15	13.15	600	0.0101	6.11

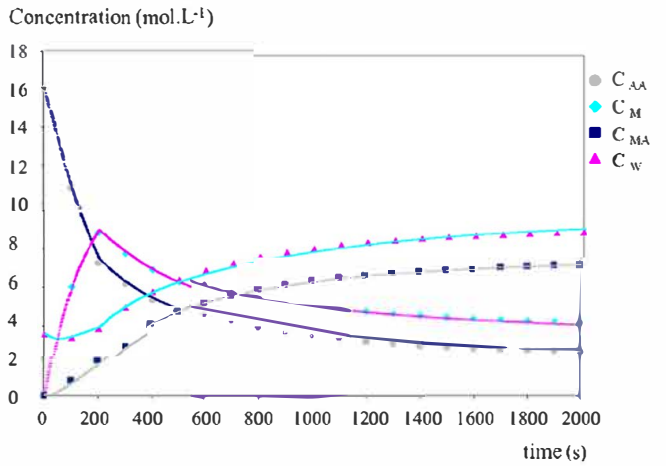


Fig. 11. Comparison of the time evolution of the different components predicted by the hybrid model (solid line) and the conventional kinetic model (points) – Fed-batch esterification at 66.55 °C with 9.14 mL of catalyst (FB1).

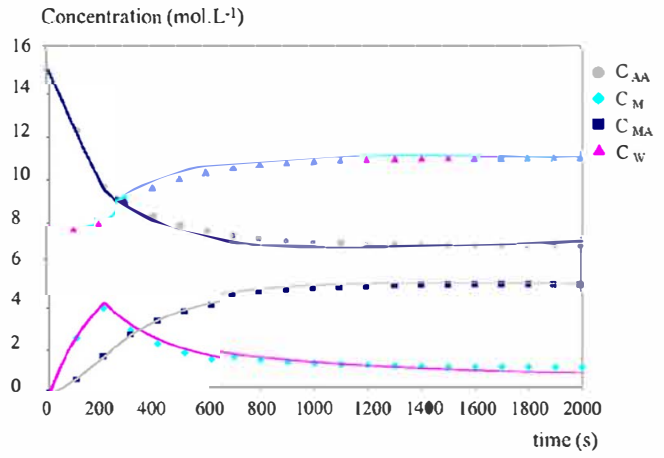


Fig. 12. Comparison of the time evolution of the different components predicted by the hybrid model (solid line) and the conventional kinetic model (points) – Fed-batch esterification at 80 °C with 7.5 mL de catalyst (FB2).

tion 2.2 (Kinetic and model parameters). A good agreement between the two types of results is observed leading to a near perfect overlap between the two curves.

The result remains almost near perfect for operating conditions at the limit (run FB2, Fig. 12) of the learning domain set in Section 4. This shows the interest of this hybrid modelling integrating a model based on neural networks that allows the generation of curves from different initial operating conditions.

The operating conditions of the FB3 run (Fig. 13), except the feeding time, were reused by increasing this feeding time to 600 s (run FB4, Fig. 14). The total number of moles of Methanol fed is constant for these two runs. Under these conditions, the results show that the change in feeding time does not affect the behavior of the hybrid model and the results perfectly reflect the time evolution of the concentrations of the different species (reagents and products).

6. Conclusion

This article has demonstrated the interest of neural networks as a tool for modelling time evolution of the different compounds during a chemical synthesis. The aim was to propose a methodol

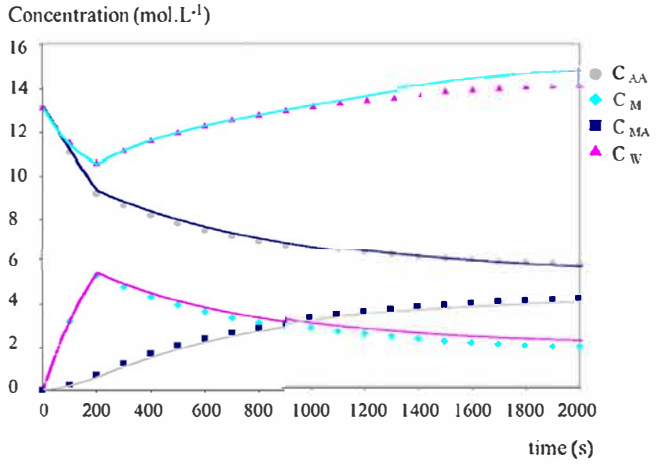


Fig. 13. Comparison of the time evolution of the different components predicted by the hybrid model (solid line) and the conventional kinetic model (points) – Fed-batch esterification at 60 °C with 8 mLde catalyst and a feeding time of 200 s (FB3).

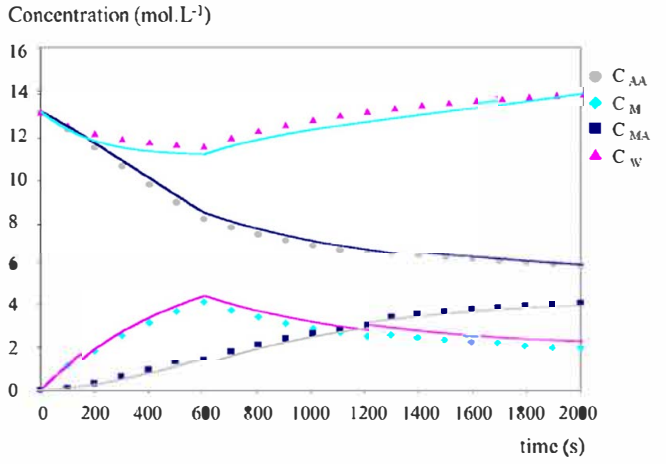


Fig. 14. Comparison of the time evolution of the different components predicted by the hybrid model (solid line) and the conventional kinetic model (points) – Fed-batch esterification at 60 °C with 8 mLof catalyst and a feeding time of 600 s (FB4).

ogy for rapidly developing a global model of the chemical transfor mation. For this, the idea was to develop a model from neural net works according to two axes: on the one hand the global architecture of the model is defined according to the constituents involved and to the influencing operating conditions (temperature, catalyst, initial concentrations...) and on the other hand the values of the model parameters are identified from a set of available data.

The objective is not to explain the physical phenomena involved but to develop a model allowing the simulation of the reaction to determine the type of reactor to be used and to optimize the oper ating conditions in relation to the technological constraints of the selected apparatus.

From the point of view of the model, an original architecture has been developed for modelling the esterification reaction of methanol by acetic acid from data obtained in isothermal batch mode. The approach was to develop an artificial neural network for the prediction of the evolution of each chemical species, then to associate these four networks in a recurrent model (called 4ANN) to obtain global predictions. The inputs of the model are the operating conditions (temperature and catalyst) as well as the concentrations of the different species at the beginning of the sampling period.

The prediction of the curves is done point by point in a recur rent mode and the only information provided to the 4ANN model are the temperature T (isothermal mode), the amount of Catalyst (sulfuric acid, Qcat) and the initial concentrations (Ci) of the four species (acetic acid, methanol, methyl acetate and water). Overall, the various studies conducted give interesting and promising results. The model (4ANN) developed makes it possible to properly estimate the evolution of the esterification reaction of methanol by acetic acid in batch mode for different operating conditions. Vali dation on experimental points confirm the validity of the 4ANN model.

The feasibility of modelling reaction kinetics by neural model has been demonstrated. The 4ANN model was developed only from data obtained in isothermal batch runs. Subsequently, from the analysis of the equations translating the material balances of a semi continuous reactor, the 4ANN model has been integrated into a hybrid simulation framework and it can be concluded that this latter is able to transpose the esterification reaction from a batch reactor to a semi batch reactor.

Once this hybrid model has been validated, it can be integrated into the process optimization of the operating conditions while considering the constraints of the apparatus and limiting the field of research to the leaming domain used to develop the neural models.

CRediT authorship contribution statement

Yessin Ammar: Investigation, Data curation, Writing original draft. **Patrick Cognet:** Conceptualization, Validation, Supervision, Writing review & editing. **Michel Cabassud:** Conceptualization, Methodology, Software, Supervision, Validation, Writing review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing finan cial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Agreda, V.H., Partin, L.R., Heise, W.H., 1990. High-purity methyl acetate via reactive distillation. Chem. Eng. Prog. February, 40–46.

Anxionnaz, Z., Cabassud, M., Gourdon, C., Tochon, P., 2010. Transposition of an exothermic reaction from a batch reactor to an intensified continuous one. Heat Transfer Eng. 31 (9), 788–797.

Balland, L., Mouhab, N., Cosmao, J.-M., Estel, L., 2002. Kinetic parameter estimation of solvent-free reactions: application to esterification of acetic anhydride by methanol. Chem. Eng. Process. 41 (Issue 5), 395–402.

Bayada, A., Lawrance, G.A., Meder, M., Molloy, K.J., 1995. ATR-IR spectroscopy for the investigation of solution reaction kinetics: hydrolysis of trimethyl phosphate. Appl. Spectrosc. 49, 1789–1792.

Bijlsma, S., Louwerse, D.J., Smilde, A.K., 1998. Rapid estimation of rate constants using SW-NIR. AIChE J. 44, 2713–2723.

Bulsari, A.B., 1995. Neural Networks for Chemical Engineers. Elsevier, Amsterdam.

Cabassud, M., Cognet, P., Garcia, V., Le Lann, M.V., Rigal, L., Casamatta, G., 2005. A global approach for the optimization of batch chemical reactors: Application to the lactic acid synthesis by the alkaline degradation of fructose. Chem. Eng. Commun. 192 (6), 758–786.

Chouai, A., Cabassud, M., Le Lann, M.V., Gourdon, C., Casamatta, G., 2000a. Use of neural networks for liquid-liquid extraction column modelling: an experimental study. Chem. Eng. Process. 39, 171–180.

Chouai, A., Cabassud, M., Le Lann, M.V., Gourdon, C., Casamatta, G., 2000b. Multivariable control of a pulsed liquid-liquid extraction column by neural networks. Neural Comput. Appl. 9, 181–189.

Delgrange-Vincent, N., Cabassud, C., Cabassud, M., Durand-Bourlier, L., Lainé, J.M., 2000. Neural networks for long-term prediction of fouling and backwash efficiency in ultrafiltration for drinking water production. Desalination 131, 353–362.

Dennis, J.E., Schnabel, R.B., 1983. Numerical Methods for Unconstrained Optimisation and Nonlinear Equations. Prentice-Hall, New Jersey.

Dirion, J.L., Cabassud, M., Le Lann, M.V., Casamatta, G., 1995. Design of a neural controller by inverse modelling. Comput. Chem. Engng. 19.

Dyson, R.M., Hazenkamp, M., Kaufmann, K., Maeder, M., Studer, M., Zilian, A., 2000. Modern tools for reaction monitoring: hard and soft modelling of non-ideal on-line acquired spectra. J. Chemom. 14, 737–750.

Elgue, S., Devatine, A., Prat, L., Cognet, P., Cabassud, M., Gourdon, C., Chopard, F., 2009. Intensification of ester production in a continuous reactor. Int. J. Chem. Reactor Eng. 7 (1), 1542–6580. https://doi.org/10.2202/1542-6580.1592. ISSN (Online).

Fakhr-Eddine, K., Cabassud, Duverneuil, P., Le Lann, M.V., Couderc, J.P., 1996. Use of neural network for LPCVD reactors modelling. Comput. Chem. Eng. 20 (Suppl.), 521–526.

Funashi, K.I., 1989. On the approximative realization of continuous mappings by neural networks. Neural Netw. 2, 183–192.

Furusjö, E., Danielsson, L.-G., 1998. A method for the determination of reaction mechanisms and rate constants from two-way spectroscopic data. Anal. Chim. Acta 373, 83–94.

Furusjö, E., Danielsson, L.-G., 2000a. Uncertainty in rate constants estimated from spectral data with baseline drift. J. Chemom. 14, 483–499.

Furusjö, E., Danielsson, L.G., 2000b. Target testing procedure for determining chemical kinetics from spectroscopic data with absorption shifts and baseline drift. Chemometrics Intellient Lab. Syst. 50, 63–73.

Furusjö, E., Svenssonb, O., Danielssona, L.G., 2003. Estimation of kinetic parameters from non-isothermal batch experiments monitored by in situ vibrational spectroscopy. Chemometrics Intelligent Lab. Syst. 66, 1–14.

Hosen, M.A., Khosravi, A., Nahavandi, S., Creighton, D., 2014. Prediction interval-based neural network modelling of polystyrene polymerization reactor – a new perspective of data-based modelling. Chem. Eng. Res. Des. 92, 2041–2051. https://doi.org/10.1016/j.cherd.2014.02.016.

Mathieu, D., Phau-Tan-Luu, R., 1981. NEMROD: New Efficient Methodology for Research using Optimal Design. L.P.R.A.I., Universite d'Aix Marseille.

Molga, E.J., 1983. Neural network approach to support modelling of chemical reactors: problems, resolution, criteria of application. Chem. Eng. Process. 42, 675–695.

Molloy, K.J., Maeder, M., Schumacher, M.M., 1999. Hard modelling of spectroscopic measurements. Appl. Non-ideal Ind. React. Syst. Chemometrics Intell. Lab. Syst. 46, 221–230.

Mozharov, S., Nordon, A., Littlejohn, D., Wiles, C., Watts, P., Dallin, P., Girkin, J.M., 2011. Improved method for kinetic studies in microreactors using flow manipulation and noninvasive Raman spectrometry. J. Am. Chem. Soc. 133, 3601–3608. https://doi.org/10.1021/ja1102234.

Mujtaba, M., Aziz, N., Hussain, M.A., 2006. Neural network based modelling and control in batch reactor. Chem. Eng. Res. Des. 84 (A8), 635–644.

Puxty, G., Maeder, M., Rhinehart, R.R., Alam, S., Moore, S., Gemperline, P.J., 2005. Modeling of batch reactions with in situ spectroscopic measurements and calorimetry. J. Chemom. 19, 329–340.

Puxty, G., Neuhold, Y.-M., Jecklin, M., Ehly, M., Gemperline, P.J., Nordon, A., Littlejohn, D., Katy Basford, J., De Cecco, M., Hungerbühler, K., 2008. Multivariate kinetic hard-modelling of spectroscopic data: a comparison of the esterification of butanol by acetic anhydride on different scales and with different instruments. Chem. Eng. Sci. 63, 4800–4809. https://doi.org/10.1016/j.ces.2008.01.020.

Rivals, I., 1995, Modélisation et commande de processus par réseaux de neurones ; application au pilotage d'un véhicule autonome, Thèse de Doctorat de l'Université Paris 6, (available at http://www.neurones.espci.fr/~rivals/).

Rizkin, B.A., Popovich, K., Hartman, R.L., 2019. Artificial Neural Network control of thermoelectrically-cooled microfluidics using computer vision based on IR thermography. Comput. Chem. Eng. 121, 584–593. https://doi.org/10.1016/j.compchemeng.2018.11.016.

Smith, H.A., 1939. Kinetics of catalyzed esterification of normal aliphatic acids in methyl alcohol. J. Am. Chem. Soc. 61, 254–260.

Stankiewicz, A., Moulijn, J.A., 2002. Process intensification. Ind. Eng. Chem. Res. 41 (8), 1920–1924.

Sunphorka, S., Chalermsinsuwan, B., Piumsomboon, P., 2017. Artificial neural network model for the prediction of kinetic parameters of biomass pyrolysis from its constituents. Fuel 193, 142–158. https://doi.org/10.1016/j.fuel.2016.12.046.

Watrous R.L., 1987. In: Proc. of IEEE First Int. Conf. Neural Networks, pp. 619–627.